

RECEIVED

MAR 17 2003



TECH CENTER 1600/2900

1600

## RAW SEQUENCE LISTING

DATE: 03/12/2003

PATENT APPLICATION: US/09/898,165B

TIME: 08:56:08

Input Set : A:\2d81552.txt

Output Set: N:\CRF4\03122003\I898165B.raw

4 <110> APPLICANT: Daniel H. Cohn  
5 Muhammad Faiyaz ul Haque  
6 Lily M. King  
7 Deborah Krakow  
9 <120> TITLE OF INVENTION: 3-Phosphoadenosine-5-Phosphosulfate  
10 (PAPS) Synthetase Proteins and Methods for Treating  
11 Osteoarthritic Disorders  
13 <130> FILE REFERENCE: 18810-81552  
15 <140> CURRENT APPLICATION NUMBER: US 09/898,165B  
16 <141> CURRENT FILING DATE: 2001-07-02  
18 <150> PRIOR APPLICATION NUMBER: 09/399,212  
19 <151> PRIOR FILING DATE: 1999-09-17  
21 <160> NUMBER OF SEQ ID NOS: 33  
23 <170> SOFTWARE: FastSEQ for Windows Version 4.0  
25 <210> SEQ ID NO: 1  
26 <211> LENGTH: 2014  
27 <212> TYPE: DNA  
28 <213> ORGANISM: Homo sapiens  
30 <400> SEQUENCE: 1

31 ctgctgccgc cgcgcgccgc gccgtccctg cgtccttcgg tctctgctcc cgggaccccg 60  
32 ctccgccgca gccagccagc atgtcgggga tcaagaagca aaagacggag aaccagcaga 120  
33 aatccaccaa tgtagtctat caggcccacc atgtgagcag gaataagaga gggcaagtgg 180  
34 ttggaacaag ggggtgggtc cgaggatgta cgtgtggct aacaggtctc tctggtgctg 240  
35 gaaaaacaac gataagtttt gccctggagg agtaccttgt ctcccatgcc atcccttgtt 300  
36 actccctgga tggggacaat gtccgtcatg gccttaacag aaatctcgga ttctctctcg 360  
37 gggacagaga ggaaaatatc cgccggattg ctgagggtgg taagctgttt gctgatgctg 420  
38 gtctggtctg cattaccagc tttattttct cattcgcaa ggatcgtgag aatgcccgca 480  
39 aaatacatga atcagcaggg ctgccattct ttgaaatatt tgtagatgca cctctaaata 540  
40 ttgttgaaag cagagacgta aaaggcctct ataaaagggc cagagctggg gagattaaag 600  
41 gatttacagg tattgattct gattatgaga aacctgaaac tctgagcgt gtgcttaaaa 660  
42 ccaatttgct cacagtgaat gactgtgtcc accaggtagt ggaacttctg caagagcaga 720  
43 acattgtacc ctatactata atcaaagata tccacgaact ctttgtgccg gaaaacaaac 780  
44 ttgaccacgt ccgagctgag gctgaaactc tcccttcatt atcaattact aagctggatc 840  
45 tccagtgggt ccagggtttg agcgaaggct gggccactcc cctcaaagggt ttcattgccg 900  
46 agaaggagta cttacagggt atgcactttg acaccctgct agatgatggc gtgatcaaca 960  
47 tgagcatccc cattgtactg cccgtctctg cagaggataa gacacggctg gaagggtgca 1020  
48 gcaagtttgt cctggcacat ggtggacgga gggtagctat cttacgagac gctgaattct 1080  
49 atgaacacag aaaagaggaa cgctgttccc gtgtttgggg gacaacatgt acaaaacacc 1140  
50 cccatatcaa aatggtgatg gaaagtgggg actggctggg tgggtggagac cttcagggtg 1200  
51 tggagaaaat aagatggaat gatgggctgg accaataccg tctgacacct ctggagctca 1260  
52 aacagaaatg taaagaaatg aatgctgatg cgggttttgc attccagttg cgcaatcctg 1320  
53 tccacaatgg tcatgccctg ttgatgcagg acacctgccg caggctccta gagaggggct 1380  
54 acaagcaccg ggtcctccta ctacaccctc tgggcgggct gaccaaggat gacgatgtgc 1440

ENTERED

## RAW SEQUENCE LISTING

DATE: 03/12/2003

PATENT APPLICATION: US/09/898,165B

TIME: 08:56:08

Input Set : A:\2d81552.txt

Output Set: N:\CRF4\03122003\I898165B.raw

```

55 ctctagactg gcggtatgaag cagcacgcgg ctgtgctcga ggaaggggtc ctggatccca 1500
56 agtcaacccat tgttgccatc ttccggtctc ccatgttata tgctggcccc acagaggtcc 1560
57 agtggcactg caggtcccg atgattgcgg gtgccaattt ctacattgtg gggagggacc 1620
58 ctgcaggaat gccccatcct gaaaccaaga aggatctgta tgaaccact catgggggca 1680
59 aggtcttgag catggcccc ggccctacct ctgtggaaat cattccattc cgagtggctg 1740
60 cctacaacaa agccaaaaaa gccatggact tctatgatcc agcaaggcac aatgagtttg 1800
61 acttcatctc aggaactcga atgaggaagc tcgcccggga aggagagaat cccccagatg 1860
62 gcttcatggc ccccaaagca tggaagggtc tgacagatta ttacagggtc ctggagaaga 1920
63 actaagcctt tgggtccaga gtttctttct gaagtgtctt ttgattacct tttctatttt 1980
64 tatgattaga tgctttgtat taaattgctt ctca 2014
66 <210> SEQ ID NO: 2
67 <211> LENGTH: 2000
68 <212> TYPE: DNA
69 <213> ORGANISM: Mus musculus
71 <400> SEQUENCE: 2
72 gtattctcaa catcagatat catgtcttgg aggaagttac ctaaactctg aagaattatc 60
73 atgtctgcaa atttcaaaat gaaccataaa agagaccagc aaaaatccac caatgtgggtc 120
74 taccaggccc atcatgtgag caggaacaag agaggacaag tggttggaac cagggggagga 180
75 ttccgaggat gtaccgtgtg gctaacagggt ctctctgggtg ctgggaaaac aaccataagc 240
76 ttgcttttgg aagagtacct tgtatctcac gccatcccat gttactccct ggatggggac 300
77 aatgtccgtc atggccttaa taagaacctg ggattctctg ccggggaccg agaagagaat 360
78 atccgccgga tcgcgagggt ggccaagctc ttgcccagc ccggcctggt ttgcatcacc 420
79 agcttttatct ctccttttgc aaaggatcgt gagaatgccc gaaaaatcca cgaatcagca 480
80 ggactcccggt tctttgagat ctttgtagat gcgcctttaa atatctgtga aagccgagac 540
81 gtaaaaggac tctacaaacg agcccagca ggagagatta aagggtttac aggcacgat 600
82 tctgactatg agaaacctga aactccagag tgtgtgctga agaccaactt gtcttcagta 660
83 agcgactgtg tgcaacagggt ggtggaactt ttgcaggagc agaacattgt accccacacc 720
84 accatcaaag gcatccacga actctttgtg ccagaaaaca aagtcgatca aatccgagct 780
85 gaggcagaga ctctcccatc actaccaatt accaagctgg atctgcagtg ggtgcagatt 840
86 ctgagtgaag gctggggcac tcccctcaaa ggctttatgc gggagaagga atacttgcaa 900
87 actctgaact tcgacactct actggacgat ggagcatca acatgagtat tccattgta 960
88 ttgcccgttt ctgcggtatga caaggcacgg ctggaagggt gcagcaaatt tgccttgatg 1020
89 tacgaaggte ggaggggtgc tctattacag gaccctgaat tctatgagca taggaaagag 1080
90 gagcgttggt ctctgtgtgt gggaacagcc actgcaaagc acccccatat caaaatggtg 1140
91 atggaaagtg gggactggct tgttgggtga gacctacagg tgctagagag aataagggtg 1200
92 gacgatgggc tggaccaata ccgccttacg cctctggaac tcaaacagaa gtgtaaagac 1260
93 atgaatgctg atgccgtgtt tgcattccag ttgcgcaatc ctgtccacaa tggatcatgcc 1320
94 ctctgatgc aggcaccccg ccgcaggctc ctggagaggg gttacaagca cccagtcctc 1380
95 ctgctccacc ctcttggggg ctggaccaag gacgatgacg tacctctgga atggaggatg 1440
96 aaacagcatg cagctgtact ggaggaaagg gtccctggatc ccaagtcaac tattgttgcc 1500
97 atctttccat ctccatgtt atacgctggt cccacagagg tccagtggca ttgcagatgc 1560
98 cggatgattg caggagccaa tttctacatt gtgggtaggg atcccgcagg aatgccccat 1620
99 cctgagacaa agaaagacct atatgaacct acccacgggg gcaaggctct gagtatggcc 1680
100 cctggcctta cctctgtgga aataattccg ttccgagtgg ctgcctacaa taaaattaaa 1740
101 aaggccatgg acttttatga tccagcaagg cagcaggagt ttgacttcat ctccaggaact 1800
102 cgcatgagga agctcgcccc ggaaggagaa gatccccag atggcttcat ggccccgaaa 1860
103 gcgtggaaag tgttgacaga ttactacagg tctctggaga agaccaacta ggtgctcctg 1920
104 gctctggctt ctccctcaag tgctctctga cgattttttt tttctatttt tgtgatttag 1980
105 ctgctctgta tccaattgca 2000

```

## RAW SEQUENCE LISTING

DATE: 03/12/2003

PATENT APPLICATION: US/09/898,165B

TIME: 08:56:08

Input Set : A:\2d81552.txt

Output Set: N:\CRF4\03122003\I898165B.raw

```

107 <210> SEQ ID NO: 3
108 <211> LENGTH: 20
109 <212> TYPE: DNA
110 <213> ORGANISM: Homo sapiens
112 <400> SEQUENCE: 3
113 tggaccaagg atgacgatgt                20
115 <210> SEQ ID NO: 4
116 <211> LENGTH: 20
117 <212> TYPE: DNA
118 <213> ORGANISM: Homo sapiens
120 <400> SEQUENCE: 4
121 cggaaagatg gcaacaatgg                20
123 <210> SEQ ID NO: 5
124 <211> LENGTH: 20
125 <212> TYPE: DNA
126 <213> ORGANISM: Homo sapiens
128 <400> SEQUENCE: 5
129 ctggtgctgg aaaaacaacg                20
131 <210> SEQ ID NO: 6
132 <211> LENGTH: 22
133 <212> TYPE: DNA
134 <213> ORGANISM: Homo sapiens
136 <400> SEQUENCE: 6
137 tgcgaatgga gaaataaagc tg            22
139 <210> SEQ ID NO: 7
140 <211> LENGTH: 615
141 <212> TYPE: PRT
142 <213> ORGANISM: Homo sapiens
144 <400> SEQUENCE: 7
145 Met Ser Gly Ile Lys Lys Gln Lys Thr Glu Asn Gln Gln Lys Ser Thr
146 1      5      10      15
147 Asn Val Val Tyr Gln Ala His His Val Ser Arg Asn Lys Arg Gly Gln
148      20      25      30
149 Val Val Gly Thr Arg Gly Gly Phe Arg Gly Cys Thr Val Trp Leu Thr
150      35      40      45
151 Gly Leu Ser Gly Ala Gly Lys Thr Thr Ile Ser Phe Ala Leu Glu Glu
152      50      55      60
153 Tyr Leu Val Ser His Ala Ile Pro Cys Tyr Ser Leu Asp Gly Asp Asn
154 65      70      75      80
155 Val Arg His Gly Leu Asn Arg Asn Leu Gly Phe Ser Pro Gly Asp Arg
156      85      90      95
157 Glu Glu Asn Ile Arg Arg Ile Ala Glu Val Ala Lys Leu Phe Ala Asp
158      100     105     110
159 Ala Gly Leu Val Cys Ile Thr Ser Phe Ile Ser Pro Phe Ala Lys Asp
160      115     120     125
161 Arg Glu Asn Ala Arg Lys Ile His Glu Ser Ala Gly Leu Pro Phe Phe
162      130     135     140
163 Glu Ile Phe Val Asp Ala Pro Leu Asn Ile Cys Glu Ser Arg Asp Val
164 145     150     155     160

```

## RAW SEQUENCE LISTING

DATE: 03/12/2003

PATENT APPLICATION: US/09/898,165B

TIME: 08:56:08

Input Set : A:\2d81552.txt

Output Set: N:\CRF4\03122003\I898165B.raw

```

165 Lys Gly Leu Tyr Lys Arg Ala Arg Ala Gly Glu Ile Lys Gly Phe Thr
166                               165                               170                               175
167 Gly Ile Asp Ser Asp Tyr Glu Lys Pro Glu Thr Pro Glu Arg Val Leu
168                               180                               185                               190
169 Lys Thr Asn Leu Ser Thr Val Ser Asp Cys Val His Gln Val Val Glu
170                               195                               200                               205
171 Leu Leu Gln Glu Gln Asn Ile Val Pro Tyr Thr Ile Ile Lys Asp Ile
172                               210                               215                               220
173 His Glu Leu Phe Val Pro Glu Asn Lys Leu Asp His Val Arg Ala Glu
174 225                               230                               235                               240
175 Ala Glu Thr Leu Pro Ser Leu Ser Ile Thr Lys Leu Asp Leu Gln Trp
176                               245                               250                               255
177 Val Gln Val Leu Ser Glu Gly Trp Ala Thr Pro Leu Lys Gly Phe Met
178                               260                               265                               270
179 Arg Glu Lys Glu Tyr Leu Gln Val Met His Phe Asp Thr Leu Leu Asp
180                               275                               280                               285
181 Asp Gly Val Ile Asn Met Ser Ile Pro Ile Val Leu Pro Val Ser Ala
182 290                               295                               300
183 Glu Asp Lys Thr Arg Leu Glu Gly Cys Ser Lys Phe Val Leu Ala His
184 305                               310                               315                               320
185 Gly Gly Arg Arg Val Ala Ile Leu Arg Asp Ala Glu Phe Tyr Glu His
186                               325                               330                               335
187 Arg Lys Glu Glu Arg Cys Ser Arg Val Trp Gly Thr Thr Cys Thr Lys
188                               340                               345                               350
189 His Pro His Ile Lys Met Val Met Glu Ser Gly Asp Trp Leu Val Gly
190                               355                               360                               365
191 Gly Asp Leu Gln Val Leu Glu Lys Ile Arg Trp Asn Asp Gly Leu Asp
192 370                               375                               380
193 Gln Tyr Arg Leu Thr Pro Leu Glu Leu Lys Gln Lys Cys Lys Glu Met
194 385                               390                               395                               400
195 Asn Ala Asp Ala Val Phe Ala Phe Gln Leu Arg Asn Pro Val His Asn
196                               405                               410                               415
197 Gly His Ala Leu Leu Met Gln Asp Thr Cys Arg Arg Leu Leu Glu Arg
198                               420                               425                               430
199 Gly Tyr Lys His Pro Val Leu Leu Leu His Pro Leu Gly Gly Trp Thr
200                               435                               440                               445
201 Lys Asp Asp Asp Val Pro Leu Asp Trp Arg Met Lys Gln His Ala Ala
202 450                               455                               460
203 Val Leu Glu Glu Gly Val Leu Asp Pro Lys Ser Thr Ile Val Ala Ile
204 465                               470                               475                               480
205 Phe Pro Ser Pro Met Leu Tyr Ala Gly Pro Thr Glu Val Gln Trp His
206                               485                               490                               495
207 Cys Arg Ser Arg Met Ile Ala Gly Ala Asn Phe Tyr Ile Val Gly Arg
208                               500                               505                               510
209 Asp Pro Ala Gly Met Pro His Pro Glu Thr Lys Lys Asp Leu Tyr Glu
210                               515                               520                               525
211 Pro Thr His Gly Gly Lys Val Leu Ser Met Ala Pro Gly Leu Thr Ser
212 530                               535                               540
213 Val Glu Ile Ile Pro Phe Arg Val Ala Ala Tyr Asn Lys Ala Lys Lys

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/898,165B

DATE: 03/12/2003

TIME: 08:56:08

Input Set : A:\2d81552.txt

Output Set: N:\CRF4\03122003\I898165B.raw

```

214 545
215 Ala Met Asp Phe Tyr Asp Pro Ala Arg His Asn Glu Phe Asp Phe Ile
216 550 555
217 Ser Gly Thr Arg Met Arg Lys Leu Ala Arg Glu Gly Glu Asn Pro Pro
218 565 570
219 Asp Gly Phe Met Ala Pro Lys Ala Trp Lys Val Leu Thr Asp Tyr Tyr
220 580 585
221 Arg Ser Glu Met Asp Lys Asn
222 595 600
225 <210> SEQ. ID NO: 8
226 <211> LENGTH: 617
227 <212> TYPE: PRT
228 <213> ORGANISM: Mus musculus
230 <400> SEQUENCE: 8
231 Met Ser Ala Asn Phe Lys Met Asn His Lys Arg Asp Gln Gln Lys Ser
232 1 5 10 15
233 Thr Asn Val Val Tyr Gln Ala His His Val Ser Arg Asn Lys Arg Gly
234 20 25 30
235 Gln Val Val Gly Thr Arg Gly Gly Phe Arg Gly Cys Thr Val Trp Leu
236 35 40 45
237 Thr Gly Leu Ser Gly Ala Gly Lys Thr Thr Ile Ser Phe Ala Leu Glu
238 50 55 60
239 Glu Tyr Leu Val Ser His Ala Ile Pro Cys Tyr Ser Leu Asp Gly Asp
240 65 70 75 80
241 Asn Val Arg His Gly Leu Asn Lys Asn Leu Gly Phe Ser Ala Gly Asp
242 85 90 95
243 Arg Glu Glu Asn Ile Arg Arg Ile Ala Glu Val Ala Lys Leu Phe Ala
244 100 105 110
245 Asp Ala Gly Leu Val Cys Ile Thr Ser Phe Ile Ser Pro Phe Ala Lys
246 115 120 125
247 Asp Arg Glu Asn Ala Arg Lys Ile His Glu Ser Ala Gly Leu Pro Phe
248 130 135 140
249 Phe Glu Ile Phe Val Asp Ala Pro Leu Asn Ile Cys Glu Ser Arg Asp
250 145 150 155 160
251 Val Lys Gly Leu Tyr Lys Arg Ala Arg Ala Gly Glu Ile Lys Gly Phe
252 165 170 175
253 Thr Gly Ile Asp Ser Asp Tyr Glu Lys Pro Glu Thr Pro Glu Cys Val
254 180 185 190
255 Leu Lys Thr Asn Leu Ser Ser Val Ser Asp Cys Val Gln Gln Val Val
256 195 200 205
257 Glu Leu Leu Gln Glu Gln Asn Ile Val Pro His Thr Thr Ile Lys Gly
258 210 215 220
259 Ile His Glu Leu Phe Val Pro Glu Asn Lys Val Asp Gln Ile Arg Ala
260 225 230 235 240
261 Glu Ala Glu Thr Leu Pro Ser Leu Pro Ile Thr Lys Leu Asp Leu Gln
262 245 250 255
263 Trp Val Gln Ile Leu Ser Glu Gly Trp Ala Thr Pro Leu Lys Gly Phe
264 260 265 270
265 Met Arg Glu Lys Glu Tyr Leu Gln Thr Leu His Phe Asp Thr Leu Leu

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/898,165B

DATE: 03/12/2003  
TIME: 08:56:09

Input Set : A:\2d81552.txt  
Output Set: N:\CRF4\03122003\I898165B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:29; N Pos. 23,305